



## Technical note

## Testing for differences in spatial distributions from individual based data

Bruce J. McAdam\*, Timothy B. Grabowski<sup>1</sup>, Guðrún Marteinsdóttir

Institute of Biology, University of Iceland, Askja, Sturlugata 7, Reykjavik 101, Iceland

## ARTICLE INFO

## Article history:

Received 29 November 2011

Received in revised form 21 February 2012

Accepted 22 February 2012

## Keywords:

Data storage tags

Individuals

Randomization

Spatial distribution

Telemetry

## ABSTRACT

Spatial distribution is increasingly studied using individual based telemetric methods in lieu of, or supplementing, surveys. Distributions may be in two- or three-dimensions, or in an abstract space such as depth-temperature space. One of the most basic questions one might address when analyzing these data is whether the distributions of two species, populations, genders, age classes, or other units are different. However due to the inherent differences between survey and telemetric approaches, it is difficult for practitioners to find a simple and easily applied approach to answer this question. In contrast to surveys, telemetry collects a large amount of data about a small number of individuals. Methods must therefore account for the random effects of individual variation in a way not otherwise necessary. We will demonstrate, for example, that tests suitable for detecting differences in distributions from survey data can give false positives (type I errors) when faced with telemetry data. This is essentially because the test treats the large number of data points as a very high  $n$ , but in fact the small number of individuals makes the  $n$  very small. As a solution, we present a test for differences in distribution based on an existing test for survey data based upon randomizing the data at the level of individuals rather than observations.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

One of the most basic questions that might be asked in ecology is whether the spatial distributions of two groups are different, be they different species, populations, genders, age classes, genotypes, or even a single population at different times. Understanding how groups differ in spatial distribution in time and space is critical for assessing historical and current patterns of habitat use and changes through time.

Syrjala (1996) devised a simple and easily applied statistical test for a difference between the spatial distributions of two populations. The test requires survey data consisting of counts of individuals or estimates of density of two populations taken from many locations throughout a landscape. The samples at each location must be independent: ideally the individuals counted will be different, as is common for many types of survey data such as fishing trawl surveys and quadrat surveys of vegetation and sessile marine organisms. This approach has been used successfully to test whether the distributions of predator and prey species differ (Wright and Begg, 1997; Swain and Wade, 2003), to evaluate the reconstructed distributions of fossilized corals (Pandolfi and

Jackson, 2001), assess spatial changes in the biomass of jellyfish over a time series (Brodeur et al., 2002), compare spatial distribution of guilds (Leach and Givnish, 1999), and as part of assessments of the effectiveness of conservation strategies (Lawler et al., 2003). However the spatial distribution of animals is increasingly being studied using individual based telemetric methods, supplementing or in lieu of more traditional survey methodologies (e.g. Rogers and White, 2007) and obtaining the distribution or habitat use of a species by making repeated observations of a relatively small number of individuals presents a problem when using the Syrjala (1996) approach. This is because the randomization method used to establish a significance value for the test statistic is not suited to repeated measures of individuals. In this case differences between individuals result in an erroneous “significant” difference between groups (type I error) because the test treats the large number of data points as a very large sample size. However in reality, the relatively small number of individuals means that the sample size and corresponding degrees of freedom is actually very small. Although such difficulties are well understood by the statistical community (e.g. Biondini et al., 1988; Aebischer et al., 1993; Otis and White, 1999) it can be difficult for practitioners to know which test to use (and how to apply it) on their data.

We present a modification of Syrjala’s test to deal with telemetric data, both as a useful tool to fisheries scientists and as an example of how to deal with detailed individual level data. In addition to dealing with the different structure to sampling, telemetry also typically produces data with continuously varying locations

\* Corresponding author.

E-mail address: [bruce@marice.is](mailto:bruce@marice.is) (B.J. McAdam).<sup>1</sup> Current address: U.S. Geological Survey, Texas Cooperative Fish & Wildlife Research Unit, Texas Tech University, Lubbock, Texas 79409-2120, USA.

rather than discrete locations, it is not clear what approach has been taken to dealing with this in the past (for example Schaefer and Fuller, 2002) and we also deal with this issue.

## 2. Theory

### 2.1. Syrjala's (1996) method

The method requires a dataset consisting of the relative densities of two groups at specific sample locations. It calculates a test statistic,  $\Psi$ , describing the magnitude of difference in distributions between the two groups, then repeatedly randomly shuffles the data between groups and calculates the same test statistic for the random groups. The final  $p$ -value is the quantile position of  $\Psi$  in the random distribution of test statistics, such that if  $\Psi$  is greater than 95% of the generated random values, then we say that the distributions are significantly different with  $p < 0.05$ . Similar randomization methods have been used in other fields such as in shape analysis of hard body parts (Smith et al., 2002). The method is limited to testing the hypothesis that there are systematic differences between the groups, and does not describe what form those differences take.

In more detail:  $(x_k, y_k)$  is the location of the  $k$ th sampling location, where  $1 \leq k \leq K$ . Sampling records a relative density (e.g. number of individuals observed) for each group,  $g \in \{1, 2\}$ , at each location,  $d_g(x_k, y_k)$  (note that our notation differs slightly from Syrjala (1996) in preparation for the modifications described later and we use  $g$  to denote a group). The densities are normalized so that the total for each group over the whole landscape is 1

$$\gamma_g(x_k, y_k) = \frac{d_g(x_k, y_k)}{\sum_{k=1}^K d_g(x_k, y_k)} \quad (1)$$

The normalized abundance over a rectangular area can be calculated by summing the density at all points inside the rectangle

$$\Gamma_g(x_0, y_0, x_1, y_1) = \sum_{\forall k: x_0 \leq x_k \leq x_1 \wedge y_0 \leq y_k \leq y_1} \gamma_g(x_k, y_k) \quad (2)$$

where  $(x_0, y_0)$  and  $(x_1, y_1)$  are the bottom-left and top right corners of the rectangle and, again,  $g$  is either 1 or 2 for the group.

The test statistic is then calculated by taking the sum of the squares of differences between  $\Gamma_1$  and  $\Gamma_2$  over a large set of rectangles. Syrjala (1996) uses every rectangle formed from one of the corners of the landscape and one of the sample points. He divides these into four sets, one for each of the corners of the sampled area. For the bottom left corner,  $(x_{\min}, y_{\min})$  of the area, this is

$$\Psi_1 = \sum_{k=1}^K (\Gamma_1(x_{\min}, x_k, y_{\min}, y_k) - \Gamma_2(x_{\min}, x_k, y_{\min}, y_k))^2 \quad (3a)$$

Similarly for the other corners of the sample space (bottom-right, top-right and top-left respectively)

$$\Psi_2 = \sum_{k=1}^K (\Gamma_1(x_k, x_{\max}, y_{\min}, y_k) - \Gamma_2(x_k, x_{\max}, y_{\min}, y_k))^2 \quad (3b)$$

$$\Psi_3 = \sum_{k=1}^K (\Gamma_1(x_k, x_{\max}, y_k, y_{\max}) - \Gamma_2(x_k, x_{\max}, y_k, y_{\max}))^2 \quad (3c)$$

$$\Psi_4 = \sum_{k=1}^K (\Gamma_1(x_{\min}, x_k, y_k, y_{\max}) - \Gamma_2(x_{\min}, x_k, y_k, y_{\max}))^2 \quad (3d)$$

And the final test statistic,  $\Psi$ , is the mean of these four components

$$\Psi = \frac{1}{4} \sum_{c=1}^4 \Psi_c \quad (4)$$

A shorter equivalent description of Syrjala's method is that it returns 1/4 of the total square-of-difference in (normalized) density over all rectangles fixed between the corners of the sampled area and survey sites.

This test statistic is then compared to the distribution of test statistics from randomized data to determine whether the actual groups differ significantly from random. The data are randomized by exchanging  $\gamma_1(x_k, y_k)$  and  $\gamma_2(x_k, y_k)$  with a probability of 1/2. The resulting randomized data meets the null-hypothesis (there is no systematic difference between the two groups) and hence the distribution of test statistics from many sets of randomized data corresponds to the distribution under the null-hypothesis. The distribution always has a long tail. In practice a large number (e.g. 1000) of randomization is performed to get a vector of values and the  $p$ -value is obtained from the rank of the true value of  $\Psi$  in this. By itself, the value of  $\Psi$  is meaningless, and not reported in results.

For the remainder of this paper, we use an alternative version of Eq. (4) that can be found in Appendix A (Eq. (A2)). The alternative version eliminates a bias towards the corners of the landscape, but the principles behind the calculation of the test statistic and  $p$ -value remain the same.

### 2.2. Extension to telemetry

Syrjala's method concerns itself with spatial survey and census data (for more methods, see Perry et al., 2002) but habitat use of populations is now often inferred from telemetric studies such as radio tracking, acoustic telemetry, data storage tags (archival tags) and passive integrated transponder (PIT) tags. Such studies always infer the distribution of each population from a relatively large number of observations of a small number of individuals. This differs from surveys that count large numbers of individuals at different locations and generally do not count the same individual more than once. Syrjala's method and others are unable to account for the fact that a small number of individuals may not accurately represent a population, and care must be taken not to mistake differences between individuals for differences between populations. For example if we monitored only two individuals of each population, then it would be wrong to infer a difference between populations, regardless of the amount of data about each individual, as the difference could come simply from differences between individuals. Specifically, we found that applying Syrjala's test to a simulated survey created from archival tagged fish from an Atlantic cod data set (Grabowski et al., 2011) reported significant differences between *any* grouping of individuals – not only ecologically meaningful groups such as sex and behavioral type, but also to meaningless groupings such as odd vs. even ID numbers. This was because the actual individual data differed significantly from data obtained by mixing observations of individuals in the Syrjala randomization method.

The solution to this problem with telemetric data is to change the way the data are randomized when establishing the distribution of the test statistic under the null hypothesis. In this case, the data from each individual must be kept together. The random test statistics must be obtained by creating random populations from the set of all individuals and not by creating random sets of observations from the set of all observations.

In a telemetry dataset, the primary piece of information is the positions that each individual was located throughout the study. For purposes of analysis, this can be expressed as  $d_{i,g}(x, y)$ , the

number of observations of individual  $i$  in group  $g$  at location  $(x, y)$ . For some studies, there will be a small number of locations in total (e.g. location of PIT tag detectors), for other studies every location will be unique (e.g. GPS tagging). The data are first normalized at an individual level, e.g. if there are a total of 100 observations of an individual then we record that each location had a “density” of  $1/100$ .

$$d'_{i,g}(x, y) = \frac{d_{i,g}(x, y)}{\sum_{\forall x, y} d_{i,g}(x, y)} \quad (5)$$

The observations are then normalized at a group level by summing the  $d'$  for each individual in group  $g$ , and dividing by the number of individuals in the group,  $|g|$  (this normalization is only strictly necessary if the groups differ in size). This results in a normalized density of observations for each population.

$$\gamma_g(x, y) = \frac{\sum_{i \in g} d'_{i,g}(x, y)}{|g|} \quad (6)$$

This can be used in Eq. (2) to find the normalized abundance over a large rectangle.

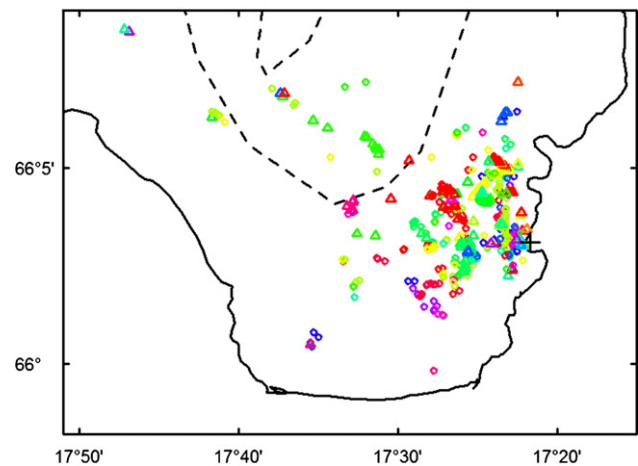
Syrjala's original test made use of the locations of sample sites when calculating the test statistic. These locations typically do not exist with telemetry data, so the test statistic is calculated using a modified version of Eq. (4) – Eq. (A2) in Appendix A.

The data are then randomized, by randomly assigning individuals to groups, and a test statistic calculated as before. An illustration of the effects of randomly assigning groups can be seen in Supplementary Fig. S1. If the experimental design ensured that there are an equal number of individuals in each group, then the random groups should also have the same number of individuals, otherwise individuals should be assigned with the probability of group membership and this may result in different sizes of randomized groups. For example, if data had 40 males and 60 females, each individual should be assigned to random group 1 with a probability of 0.4, which will result in groups of variable size averaging 40 and 60. This more accurately models the design of the experiment in the simulation.

A simple implementation of the method can be found in Supplementary information.

### 3. Methods and materials

We used three datasets to evaluate the performance of our modifications to the Syrjala (1996) test. Our first dataset was generated from land-based observations of cetaceans in Skjálfandi Bay, northern Iceland, taken on 40 days between February and September 2009 (E.E. Magnúsdóttir, University of Iceland, unpublished data). An observer stationed on a cliff-top used a theodolite equipped with a data recorder to track visible cetaceans. Observations were translated into local co-ordinates (Lerczak and Hobbs, 1998) and grouped into tracks known to be the same individual or group of individuals. Most tracks were identified with the species being observed, although it was not possible to identify the smallest cetaceans to species level. From the cetacean data, it was clear from the map that some species had completely non-overlapping distributions. For the smallest cetaceans, minke whales (*Balaenoptera acutorostrata*) and dolphins and porpoises (*Tursiops truncatus* and *Phocoena phocoena*: in a single category as it was not possible for the observer to identify these to species level), the distributions appeared to be similar (Fig. 1). These groups have similar sightability so any apparent differences in distribution are unlikely to arise from differences in detection probability alone. Therefore we wished to test whether there were in fact statistical differences between distributions of minke whales and the smaller cetaceans. The data consist of 467 point observations grouped into

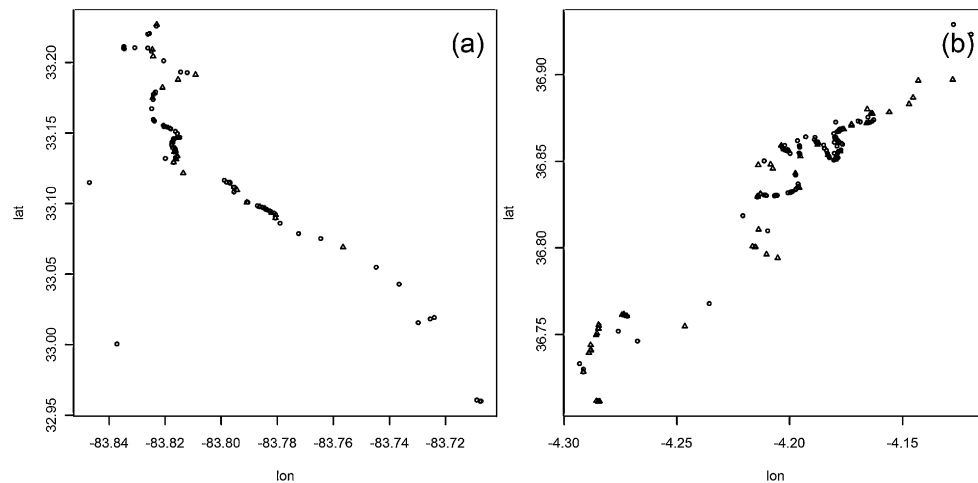


**Fig. 1.** Map of sightings of cetaceans in Skjálfandi bay, northern Iceland. This shows the typical sort of data that the test may be applied to. Solid line: coastline, dashed line: depth contours at 100 m intervals, circles: Minke whales, triangles: dolphins or porpoises, different colors are different individuals, cross: location of observer.

36 individual tracks. These points provide the values of  $d_g(x_k, y_k)$  required for Eq. (1) (observed density being 1 if an individual in group  $g$  was observed at point  $(x_k, y_k)$ , and 0 otherwise). The test statistic was calculated using  $J = 1000$  rectangles, and the  $p$ -value was calculated using 1000 random groupings.

Secondly, we used data from two radio telemetry studies of robust redhorse *Moxostoma robustum*, an imperiled North American riverine fish (Grabowski and Isely, 2006; Grabowski and Jennings, 2009). One study was conducted on the Ocmulgee River, Georgia in 2006–2007 using individuals transplanted from two refugial populations (Grabowski and Jennings, 2009): the Broad River, Georgia – a relatively fast flowing piedmont river and the Ogeechee River, Georgia – a slower flowing, coastal plain river. The data set contained 30 individuals and a total of 1207 observations (providing values for  $d_g(x_k, y_k)$  as in the cetacean example). In contrast to the cetacean observations, rather than being spread over a wide geographic region (sea surface), the observations are restricted to a linear feature, although the width of the river offers the possibility of spatial differences in this dimension (e.g. by different depth preference placing one group further from the banks). Hence this data set challenges the statistical test to identify differences in distribution of a different nature than the previous example. We compared the distributions of individuals originating from Broad River and Ogeechee River (Fig. 2a), expecting to see a difference in distribution, and also the difference in distribution of males and females (expecting no difference).

The other radio-tracking study was performed on the Savannah River, Georgia and South Carolina from 2002 to 2005 (Grabowski and Isely, 2006). There were several flood pulses, both natural and anthropogenic, in early 2003 on the Savannah River, and we used these data to test whether the distribution of the tagged fish was different before and after these events (Fig. 2b). The data set features 17 individuals (631 observations) and similarly to the second example, the observations follow a linear feature. The Savannah River study differs from the Skjálfandi Bay cetacean and the Ocmulgee River datasets as we wish to compare the same individuals at different times, rather than different individuals at the same time, a typical question in ecological studies of the temporal variability of spatial distributions. The data were divided by both individual and time period – essentially every individual provided two sets of observations each in a different group: the individual from the first time period, and the individual in the second. The test statistic was computed as the difference between the two time periods. To generate each set of random ‘null-hypothesis’ data, we then randomly



**Fig. 2.** Radio telemetry data for *M. robustum* at (a) Ocmulgee River (triangles are fish originally from Broad River and circles from Ogeechee River) and (b) Savannah River (circle before and triangles after flooding events). Both figures show only a subset of the geographic range of the data.

exchanged the data from the two time periods with a probability of 1/2. This resulted in two sets of data each drawn equally from the two periods that could be compared to compute a null-hypothesis test statistic. As a check on this method, the data were also divided into two periods for which we expected no significant difference in distribution: summer versus winter (June–August vs. December–February) over the whole time period (June 2002–May 2005).

#### 4. Results and discussion

We obtained a  $p$ -value of 0.009 when testing for differences in distribution between minke whales ( $n = 17$  tracks, 313 observations) and dolphins/porpoises ( $n = 19$  tracks, 154 observations), suggesting that minke whales were distributed differently from the smaller cetaceans in Skjálfandi Bay. The test allowed us to reject the null hypothesis of identical distributions even though the map of species distribution was difficult to interpret by eye.

Two tests were performed with the data from the Ocmulgee River. Comparing fish originating from the Broad River ( $n = 10$  individuals, 400 observations) against those from the Ogeechee River ( $n = 20$  individuals, 807 observations) yielded  $p = 0.033$ . Comparing the two sexes ( $n_{\text{male}} = 23$  individuals, 790 observations,  $n_{\text{female}} = 7$ , 417) yielded  $p = 0.114$ . The distribution of individuals from the two refugial populations were significantly different as was suggested by plots of the data showing the Broad River fish tended to be clustered upstream of the Ogeechee River fish (Grabowski and Jennings, 2009). There were no differences in the distributions of the two sexes. The test returned a significant result for the grouping that we expected to be ecologically meaningful, and accepted the null hypothesis for the other grouping. This demonstrates that the modified randomization procedure does not mistake individual variation for systematic differences with groups, as the raw observation based randomization of the original Syrjala test does.

Comparing the distribution of robust redbreast during pre- and post-flood periods on the Savannah river, we found significant differences ( $p < 0.001$ ,  $n = 17$  individuals, 631 observations), but comparing summer against winter distributions there was no significant difference ( $p = 0.44$ ,  $n = 17$  individuals, 887 observations). By modifying the way that the data was divided into groups and randomized, it was also possible to compare the same individuals over two time periods. We found that robust redbreast exhibited different distributions before and after flood events in the Savannah River. However, care should be taken to clearly define the time

periods used, as our test as currently constructed is not capable of integrating across a continuous time line.

#### 5. Conclusions

The Syrjala (1996) test was modified in the way that it samples rectangles from landscape being studied. This change was intended to reduce the bias of the test towards the center of the landscape and to allow it to be used with telemetry data that is not taken from fixed sample locations. For telemetry data consisting of large numbers of observations of small numbers of individuals, we proposed a new randomization method to generate the distribution of test statistics required to calculate a  $p$ -value. The test performed well both with examples where we expected the null-hypothesis to be rejected, and those in which no significant differences were expected. The test was also applicable both to data that had a wide spatial distribution and data that followed a linear feature.

Furthermore, we suggest that the test might be extended to three or more dimensions by sampling cuboids and ultimately, this might be the way to better incorporate a temporal aspect to the analysis. Incorporation of a third dimension will be relevant to data such as data storage tags (DSTs) for fish that record swimming depth (e.g. data as in Neat et al., 2005). The test may also be applied to abstract spaces, such as niche space analysis (Grabowski et al., 2011).

The test described here is an illustration of the difficulties of extrapolating population-level conclusions from data obtained from a few individuals – no matter how much information is collected from each individual. Although it is tempting to simulate survey-type data from telemetry, this must be avoided as individual variation almost always leads us to false conclusions. The solution, though, is essentially simple: calculate a test statistic from the real data and then compare it with values obtained by randomizing the real data so that it meets the null hypothesis. By making the test in this way, it is clear exactly what null-hypothesis is being tested and we can be sure that the results we obtain come directly from the data. This approach is likely to have wider applications in analyzing telemetry data than just in spatial distribution.

#### Acknowledgments

We are grateful to Jose M. Blanco-Moreno, University of Barcelona and Marcelino de la Cruz Rot, Universidad Politécnica de Madrid for providing an implementation of Syrjala's method in



R (the *ecspa* package), and to Stephen E. Syrjala at the National Oceanic and Atmospheric Administration, United States Department of Commerce for communication about his method prior to our work commencing. Members of the MARICE group and statistics department at the University of Iceland provided comments and suggestions for an early draft of this manuscript. Support for this project was provided by the University of Iceland Research Fund; the Icelandic Research Fund; and the Fisheries Project Fund of the Icelandic Ministry of Fisheries and Agriculture. Whale observation data was provided courtesy of Edda Elísabet Magnúsdóttir at The University of Iceland and was supported by the University of Iceland Research Fund. Funding for the collection of the Ocmulgee River robust redhorse telemetry dataset provided by Georgia Power (Project number: 20-21-RR272-077). Funding for the collection of the Savannah River robust redhorse telemetry dataset was provided by the U.S. Army Corps of Engineers, Environmental Resources Branch.

## Appendix A. Modifying the calculation for continuous locations

Syrjala's (1996) method requires data from a fixed number of specific sample locations and these locations are used to calculate the test statistic. In order to deal with the continuously varying locations from telemetry we must modify the calculation, and aim for theoretic ideal that the test statistic is the mean difference in average density over all rectangles centered on all points on the landscape. Clearly it is not computationally practical to calculate this ideal value, therefore we sample  $J$  random rectangles from the landscape, compute the difference in average density ( $\Gamma_1, \Gamma_2$ ) over each rectangle, and find the sum of squares of these differences ( $\Psi$ ). The distribution of test statistics for randomized data and the  $p$ -values are then calculated as before.

Each point in the sampled space must have equal probability of appearing in a randomly selected rectangle. There are  $J$  rectangles, each described by its  $x$  and  $y$  limits, i.e. rectangle number  $j$  is described by  $x_{j,0}, x_{j,1}, y_{j,0}$  and  $y_{j,1}$ . If the  $x$  coordinates in the data range from  $x_{\min}$  to  $x_{\max}$ , then the left hand edge of each rectangle is randomly sampled from a uniform distribution

$$x_{j,0} \sim U(x_{\min} - (x_{\max} - x_{\min}), x_{\max}) \quad (\text{A1a})$$

and the width of the rectangle is sampled from

$$w_j \sim U(0, 2(x_{\max} - x_{\min})) \quad (\text{A1b})$$

to give the right-hand edge

$$x_{j,1} = x_{j,0} + w_j \quad (\text{A1c})$$

similarly for  $y$ .

$$y_{j,0} \sim U(y_{\min} - (y_{\max} - y_{\min}), y_{\max}) \quad (\text{A1d})$$

$$h_j \sim U(0, 2(y_{\max} - y_{\min})) \quad (\text{A1e})$$

$$y_{j,1} = y_{j,0} + h_j \quad (\text{A1f})$$

This results in some rectangles that contain the whole of the sampled area, and some too small to contain any sample sites but the difference in normalized abundance over these areas is zero, and they do not affect subsequent calculations.

Eqs. (3a)–(3d) and (4) are replaced by the sum of squares of differences over all  $J$  rectangles

$$\Psi = \sum_{j=1}^J (\Gamma_1(x_{j,0}, x_{j,1}, y_{j,0}, y_{j,1}) - \Gamma_2(x_{j,0}, x_{j,1}, y_{j,0}, y_{j,1}))^2 \quad (\text{A2})$$

To find the  $p$ -value, the data are randomized and set of randomized  $\Psi$  are calculated as before but using Eq. (6).

We now demonstrate that anchoring the rectangles used in the original calculation of  $\Psi$  to the corners of the sampled space leads to different  $p$ -values depending on whether the difference in distributions is closer to the center or corner of this area. The corner sensitivity problem arises because sample sites close to the corners of the grid contribute to both large and small rectangles in the calculations of  $\Gamma_i$  and  $\Psi$ , but points close to the center only contribute to large rectangles. This makes the original method sensitive not only to differences in distribution, but also to the details of the sampling regime and in some cases can result in a type II error when the key part of the distribution lies in the wrong part of the sample grid. We perform a simulation to demonstrate that the new calculation presented here reduces the chance of type II errors, and hence suggest it should be used even where data are obtained from discrete locations.

### A.1. Method

We simulated the distribution of two species over a landscape. The background density of both species was  $1 \pm 1$  (mean  $\pm$  sd) across the whole landscape. Each species had a hotspot of radius 500 m and density  $3 \pm 1$ . The center of species A's hotspot was 400 m west of the center of species B's. We then simulated two different sampling regimes. Both sampling regimes sampled a rectangular grid at 100 m intervals over a  $4 \text{ km} \times 4 \text{ km}$  square (1600 points). The first sampling regime had the two hotspots in the center of the sampling grid, the second had the hotspots closer to the south-west corner (center of A's hotspot was 1 km north and 800 m east of the SW corner). The random noise at each location was the same for both sampling regimes. For both sampling regimes the whole of both hotspots was inside the grid and a correctly functioning test should yield similar  $p$  values for the test that the distributions are identical, but we hypothesized that the  $p$  value when the hotspots are near the center of the sample grid will be lower (more significant). We used an implementation of Syrjala's test using 1000 randomizations of the data to calculate the null-hypothesis test statistic distribution (R, version 2.9, R Development Core Team, R Foundation for Statistical Computing, Vienna Austria 2009; R *ecspa* package, Jose M. Blanco-Moreno, University of Barcelona and Marcelino de la Cruz Rot, Universidad Politécnica de Madrid 2008).

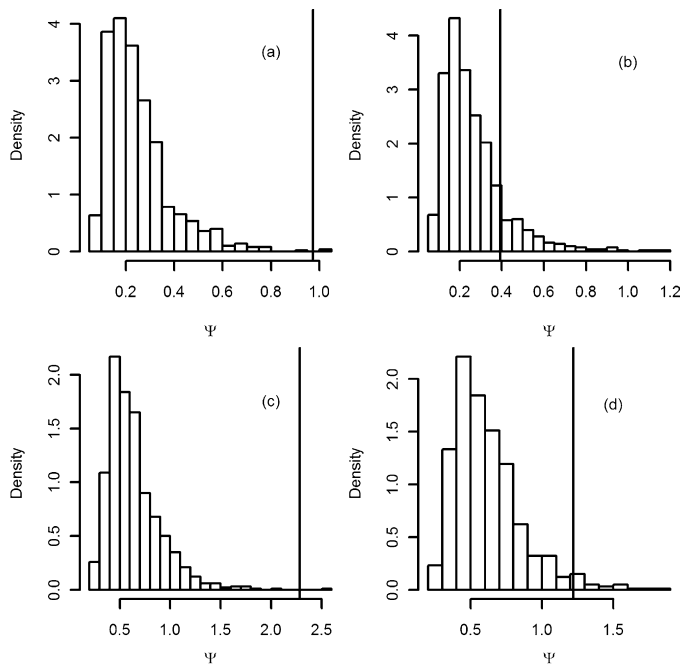
The same data (both sampling regimes) were used with the modified algorithm (Eq. (6)). The algorithm was implemented in R. The hypothesis that the modified version is less sensitive to the location of the sample grid will be confirmed if the  $p$  values for the two sampling regimes are closer together than for the original method. The test statistic was calculated using 6400 rectangles (same as the unmodified algorithm), and the null hypothesis distribution was obtained from 1000 randomizations.

The new algorithm also was used on data that met the null hypothesis (data as for the corner sensitivity test, but with no hotspot). The test should yield a non-significant  $p$  value.

### A.2. Results and discussion

Running Syrjala's test on the simulated landscape yielded a significant difference ( $p=0.002$ ) for the sampling regime with hotspots in the center of the grid, but indicated no difference in distribution ( $p=0.14$ ) when the hotspots positioned in the corner of the sampling grid. However, using the modified test with randomly sampled rectangles yielded different  $p$ -values ( $p=0.001$  when anomaly is in center of grid; and  $p=0.03$  when it is in corner). Fig. A1 shows distribution of test statistics for the four tests.

For the simulated landscape, the modified test changed the non-significant result (obtained when the anomaly was close to



**Fig. A1.** Results of the test for corner-sensitivity using simulated data. This illustrates the long tail of the null-hypothesis distribution. Histogram shows distribution of test statistics for the randomized data (null-hypothesis), vertical line shows value for actual data. (a) The original test with data from a sampling grid centered on the anomaly, (b) original test with anomaly at corner of grid, (c) modified test with anomaly in center, and (d) modified test with anomaly at corner. Note that (a, b and d) show significant  $p$ -values, as the true value lies to the far right (top 5th percentile) of the distribution, whereas the original test fails to give a significant  $p$ -value when the anomaly is at the corner of the sampled area in (c).

a corner) to into a statistically significant result with  $p < 0.05$ , although the  $p$ -values were not exactly equal for the two sampling grids so edge effects do still exist. For a simulated landscape in which there was not difference between groups, both the original and modified tests accepted the null hypothesis. Hence we recommend this as an improvement in the method for survey data even where data are obtained from discrete locations. The result demonstrates that whichever test is used, it is important that the experimental design tries to center sampling grids on the expected location of any anomalies, but our modified test seems to be robust to variation in sampling design.

## Appendix B. Supplementary information

Supplementary information associated with this article can be found, in the online version, at doi:10.1016/j.fishres.2012.02.024.

## References

- Aebischer, N.J., Robertson, P.A., Kenward, R.E., 1993. Compositional analysis of habitat use from animal radio-tracking data. *Ecology* 74, 1313–1325.
- Biondini, M.E., Mielke Jr., P.W., Berry, K.J., 1988. Data-dependent permutation techniques for the analysis of ecological data. *Vegetatio* 75, 161–168.
- Brodeur, R.D., Sugisaki, H., Hunt Jr., G.L., 2002. Increases in jellyfish biomass in the Bering Sea: implications for the ecosystem. *Mar. Ecol. Prog. Ser.* 233, 89–103.
- Grabowski, T.B., Isely, J.J., 2006. Seasonal and diel movements and habitat use of robust redbreast in the lower Savannah River, Georgia and South Carolina. *Trans. Am. Fish. Soc.* 135, 1145–1155.
- Grabowski, T.B., Jennings, C.A., 2009. Post-release movements and habitat use of robust redbreast transplanted to the Ocmulgee River, Georgia. *Aquat. Conserv.* 19, 170–177.
- Grabowski, T.B., Thorsteinsson, V., McAdam, B.J., Marteinsdóttir, G., 2011. Sympatric divergence in Atlantic cod. *PLoS ONE* 6 (3), e17528.
- Lawler, J.J., White, D., Sifneos, J.C., Master, L.L., 2003. Rare species and the use of indicator groups for conservation planning. *Conserv. Biol.* 17, 875–882.
- Leach, M.K., Givnish, T.J., 1999. Gradients in the composition, structure, and diversity of remnant oak savannahs in southern Wisconsin. *Ecol. Monogr.* 69, 353–374.
- Lerczak, J.A., Hobbs, R.C., 1998. Calculating sighting distances from angular readings during shipboard, aerial and shore-based marine mammal surveys. *Mar. Mamm. Sci.* 14, 590–598.
- Neat, F.C., Wright, P.J., Zuur, A.F., Gibb, I.M., Gibb, F.M., Tulett, D., Righton, D.A., Turner, R.J., 2005. Residency and depth movements of a coastal group of Atlantic cod (*Gadus morhua* L.). *Mar. Biol.* 148, 643–654.
- Otis, D.L., White, G.C., 1999. Autocorrelation of location estimates and the analysis of radiotracking data. *J. Wildl. Manag.* 63, 1039–1044.
- Pandolfi, J.M., Jackson, J.B.C., 2001. Community structure of Pleistocene coral reefs of Curacao, Netherlands Antilles. *Ecol. Monogr.* 71, 49–67.
- Perry, J.N., Liebhold, A.M., Rosenberg, M.S., Dungan, J., Miriti, M., Jakomulska, A., Citron-Pousty, S., 2002. Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data. *Ecography* 25, 578–600.
- Rogers, K.B., White, G.C., 2007. Analysis of movement and habitat use from telemetry data. In: Guy, C.S., Brown, M.L. (Eds.), *Analysis and Interpretation of Freshwater Fisheries Data*. American Fisheries Society, Bethesda, Maryland, pp. 625–676.
- Schaefer, K.M., Fuller, D.W., 2002. Movements, behavior, and habitat selection of big-eye tuna (*Thunnus obesus*) in the eastern equatorial Pacific, ascertained through archival tags. *Fish. Bull.* 100, 765–788.
- Smith, P.J., Robertson, S.G., Horn, P.L., Bull, B., Anderson, O.F., Stanton, B.R., Oke, C.S., 2002. Multiple techniques for determining stock relationships between orange roughy, *Hoplostethus atlanticus*, fisheries in the eastern Tasman Sea. *Fish. Res.* 58, 119–140.
- Swain, D.P., Wade, E.J., 2003. Spatial distribution of catch and effort in a fishery for snow crab (*Chionoecetes opilio*): tests of predictions of the ideal free distribution. *Can. J. Fish. Aquat. Sci.* 60, 897–909.
- Syrjala, S.E., 1996. A statistical test for a difference between the spatial distributions of two populations. *Ecology* 77, 75–80.
- Wright, P.J., Begg, G.S., 1997. A spatial comparison of common guillemots and sandeels in Scottish waters. *ICES J. Mar. Sci.* 54, 578–592.